



IFWO

RAW SEQUENCE LISTING

DATE: 08/20/2004

PATENT APPLICATION: US/10/817,418

TIME: 10:44:03

Input Set : A:\38002071.app

Output Set : N:\CRF4\08202004\J817418.raw

ENTERED

3 <110> APPLICANT: YANG, JIANXIN
 5 <120> TITLE OF INVENTION: GENE AMPLIFICATION AND OVEREXPRESSION IN CANCER
 7 <130> FILE REFERENCE: 38002-0071
 9 <140> CURRENT APPLICATION NUMBER: 10/817,418
 10 <141> CURRENT FILING DATE: 2004-04-05
 12 <150> PRIOR APPLICATION NUMBER: 60/462,702
 13 <151> PRIOR FILING DATE: 2003-04-15
 15 <160> NUMBER OF SEQ ID NOS: 23
 17 <170> SOFTWARE: PatentIn Ver. 3.2
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 3001
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Homo sapiens
 24 <400> SEQUENCE: 1
 25 tgagtaaadc gatacatcat acgcgcgcctc ctctggccgc ccctccctcc gacgacggg 60
 26 gaccctggcg ggcggcagga ggacatggcc agcgacgccg tgcagagtga gcctcgagc 120
 27 tggccctgc tagagcagct gggcctggcc ggggcagacc tggcgcccc cggggtacag 180
 28 cagcagctgg agctggagcg ggagcggctg cggcgggaaa tccgcaagga gctgaagctg 240
 29 aaggaggggtg ctgagaacct ggcggcgggc accactgacc tgggcccgcag cctgggcccc 300
 30 gtagagctgc tgctgcgggg ctccctgcgc cgctcgacc tgctgcacca gcagctgcag 360
 31 gagctgcacg cccacgtggt gcttcccgac cggcgggcca cccacgatgg ccccgagtc 420
 32 cctgggtgcg gtggccccac ctgctcgccc accaacctga gccgcgtggc gggcctggag 480
 33 aagcagttgg ccattgagct gaaggtgaag cagggggcgg agaacatgat ccagacctac 540
 34 agcaatggca gcaccaagga ccggaagctg ctgctgacag cccagcagat gttgcaggac 600
 35 agtaagacca agattgacat catccgcag caactccgcc gggcgctgca ggccgaccag 660
 36 ctggagaacc aggcagcccc ggatgacacc caagggagtc ctgacctggg ggctgtggag 720
 37 ctgcgcacg aagagctgcg gcaccacttc cgagtggagc acgcggtggc cgaggggtgc 780
 38 aagaacgtac tgcgcctgct cagcgctgcc aaggcccccg accgcaaggc agtcagcgag 840
 39 gccaggaga aattgacaga atccaaccag aagctggggc tgctgcggga ggctctggag 900
 40 cggagacttg gggagctgcc cgccgaccac cccaaggggc ggctgctgcg agaagagctc 960
 41 gctgcggcct cctccgctgc ctccagcacc cgccctggcg ggccctttcc cgccacgcac 1020
 42 tacagacacc tgtgcaagcc cgcgcgcctc acagggaccc tggaggtacg agtgggtggc 1080
 43 tgcagagacc tcccagagac catcccgtgg aaccctaccc cctcaatggg gggacctggg 1140
 44 accccagaca gccgcccccc ctccctgagc cgcccagccc ggggccttta cagccgaagc 1200
 45 ggaagcctca gtggccggag cagcctcaaa gcagaagccg agaacaccag tgaagtcagc 1260
 46 actgtgctta agctggataa cacagtgggt gggcagacgt cttggaagcc atgtggcccc 1320
 47 aatgcctggg accagagctt cactctggag ctggaaaggg cacgggaact ggagttggct 1380
 48 gtgttctggc gggaccagcg gggcctgtgt gccctcaat tcctgaagtt ggaggatttc 1440
 49 ttggacaatg agaggcatga ggtgcagctg gacatggaac cccagggctg cctggtggct 1500
 50 gaggtcacct tccgcaaccc tgtcattgag aggatccctc ggctccgacg gcagaagaaa 1560
 51 attttctcca agcagcaagg gaaggcgctc cagcgtgcta ggcagatgaa catcgatgct 1620
 52 gccacgtggg tgcggctgct ccggaggctc atccccaatg ccacgggcac aggcaccttt 1680
 53 agccctgggg cttctccagg atccgaggcc cggaccacgg gtgacatatc ggtggagaag 1740

RAW SEQUENCE LISTING

DATE: 08/20/2004

PATENT APPLICATION: US/10/817,418

TIME: 10:44:03

Input Set : A:\38002071.app

Output Set: N:\CRF4\08202004\J817418.raw

```

54 ctgaacctcg gcactgactc ggacagctca cctcagaaga gctcgcggga tcctccttcc 1800
55 agcccatcga gcttgagctc ccccatccag gaatccactg ctcccgagct gccttcggag 1860
56 acccaggaga ccccgagccc cgccctgtgc agccctctga ggaagtcacc tctgacctc 1920
57 gaagatttca agttcctggc ggtgctgggc cggggctcatt ttgggaagggt gctcctctcc 1980
58 gaattccggc ccagtgggga gctgttcgcc atcaaggctc tgaagaaagg ggacattgtg 2040
59 gcccagagacg aggtggagag cctgatgtgt gagaagcgga tattggcggc agtgaccagt 2100
60 gcgggacacc ccttcctggt gaacctcttc ggctgtttcc agacaccgga gcacgtgtgc 2160
61 ttctgatagg agtactcggc cgggtggggac ctgatgctgc acatccacag cgacgtgttc 2220
62 tctgagcccc gtgccatctt ttattccgcc tgcgtggtgc tgggcctaca gtttcttcac 2280
63 gaacacaaga tcgtctacag ggacctgaag ttggacaatt tgctcctgga caccgagggc 2340
64 tacgtcaaga tcgcagactt tggcctctgc aaggaggggga tgggctatgg ggaccggacc 2400
65 agcacattct gtgggacccc ggagttcctg gccctgagg tgctgacgga cacgtcgtac 2460
66 acgcgagctg tggactgggt gggactgggt gtgctgctct acgagatgct gggtggcgag 2520
67 tccccattcc caggggatga tgaggaggag gtcttcgaca gcatcgtcaa cgacgaggtt 2580
68 cgctaccccc gcttcctgtc ggccgaagcc atcggcatca tgagaaggct gcttcggagg 2640
69 aaccagagc ggaggctggg atctagcgag agagatgcag aagatgtgaa gaaacagccc 2700
70 ttcttcagga ctctgggctg ggaagccctg ttggcccggc gctgccacc gccctttgtg 2760
71 cccacgctgt ccggccgcac cgacgtcagc aacttcgacg aggagttcac cggggaggcc 2820
72 cccacactga gcccgccccg cgacgcgcgg cccctcacag ccgcggagca ggcagccttc 2880
73 ctggacttcg acttcgtggc cgggggctgc tagcccccct cctgcccct gccctgccc 2940
74 ctgcccagaga gctcttagtt ttaaaaagg cctttgggat ttgccgaaa aaaaaaaaaa 3000
75 a 3001

```

78 <210> SEQ ID NO: 2

79 <211> LENGTH: 942

80 <212> TYPE: PRT

81 <213> ORGANISM: Homo sapiens

83 <400> SEQUENCE: 2

```

84 Met Ala Ser Asp Ala Val Gln Ser Glu Pro Arg Ser Trp Ser Leu Leu
85 1 5 10 15
87 Glu Gln Leu Gly Leu Ala Gly Ala Asp Leu Ala Ala Pro Gly Val Gln
88 20 25 30
90 Gln Gln Leu Glu Leu Glu Arg Glu Arg Leu Arg Arg Glu Ile Arg Lys
91 35 40 45
93 Glu Leu Lys Leu Lys Glu Gly Ala Glu Asn Leu Arg Arg Ala Thr Thr
94 50 55 60
96 Asp Leu Gly Arg Ser Leu Gly Pro Val Glu Leu Leu Leu Arg Gly Ser
97 65 70 75 80
99 Ser Arg Arg Leu Asp Leu Leu His Gln Gln Leu Gln Glu Leu His Ala
100 85 90 95
102 His Val Val Leu Pro Asp Pro Ala Ala Thr His Asp Gly Pro Gln Ser
103 100 105 110
105 Pro Gly Ala Gly Gly Pro Thr Cys Ser Ala Thr Asn Leu Ser Arg Val
106 115 120 125
108 Ala Gly Leu Glu Lys Gln Leu Ala Ile Glu Leu Lys Val Lys Gln Gly
109 130 135 140
111 Ala Glu Asn Met Ile Gln Thr Tyr Ser Asn Gly Ser Thr Lys Asp Arg
112 145 150 155 160
114 Lys Leu Leu Leu Thr Ala Gln Gln Met Leu Gln Asp Ser Lys Thr Lys
115 165 170 175

```

RAW SEQUENCE LISTING

DATE: 08/20/2004

PATENT APPLICATION: US/10/817,418

TIME: 10:44:03

Input Set : A:\38002071.app

Output Set: N:\CRF4\08202004\J817418.raw

```

117 Ile Asp Ile Ile Arg Met Gln Leu Arg Arg Ala Leu Gln Ala Asp Gln
118           180           185           190
120 Leu Glu Asn Gln Ala Ala Pro Asp Asp Thr Gln Gly Ser Pro Asp Leu
121           195           200           205
123 Gly Ala Val Glu Leu Arg Ile Glu Glu Leu Arg His His Phe Arg Val
124           210           215           220
126 Glu His Ala Val Ala Glu Gly Ala Lys Asn Val Leu Arg Leu Leu Ser
127 225           230           235           240
129 Ala Ala Lys Ala Pro Asp Arg Lys Ala Val Ser Glu Ala Gln Glu Lys
130           245           250           255
132 Leu Thr Glu Ser Asn Gln Lys Leu Gly Leu Leu Arg Glu Ala Leu Glu
133           260           265           270
135 Arg Arg Leu Gly Glu Leu Pro Ala Asp His Pro Lys Gly Arg Leu Leu
136           275           280           285
138 Arg Glu Glu Leu Ala Ala Ala Ser Ser Ala Ala Phe Ser Thr Arg Leu
139           290           295           300
141 Ala Gly Pro Phe Pro Ala Thr His Tyr Ser Thr Leu Cys Lys Pro Ala
142 305           310           315           320
144 Pro Leu Thr Gly Thr Leu Glu Val Arg Val Val Gly Cys Arg Asp Leu
145           325           330           335
147 Pro Glu Thr Ile Pro Trp Asn Pro Thr Pro Ser Met Gly Gly Pro Gly
148           340           345           350
150 Thr Pro Asp Ser Arg Pro Pro Phe Leu Ser Arg Pro Ala Arg Gly Leu
151           355           360           365
153 Tyr Ser Arg Ser Gly Ser Leu Ser Gly Arg Ser Ser Leu Lys Ala Glu
154           370           375           380
156 Ala Glu Asn Thr Ser Glu Val Ser Thr Val Leu Lys Leu Asp Asn Thr
157 385           390           395           400
159 Val Val Gly Gln Thr Ser Trp Lys Pro Cys Gly Pro Asn Ala Trp Asp
160           405           410           415
162 Gln Ser Phe Thr Leu Glu Leu Glu Arg Ala Arg Glu Leu Glu Leu Ala
163           420           425           430
165 Val Phe Trp Arg Asp Gln Arg Gly Leu Cys Ala Leu Lys Phe Leu Lys
166           435           440           445
168 Leu Glu Asp Phe Leu Asp Asn Glu Arg His Glu Val Gln Leu Asp Met
169           450           455           460
171 Glu Pro Gln Gly Cys Leu Val Ala Glu Val Thr Phe Arg Asn Pro Val
172 465           470           475           480
174 Ile Glu Arg Ile Pro Arg Leu Arg Arg Gln Lys Lys Ile Phe Ser Lys
175           485           490           495
177 Gln Gln Gly Lys Ala Phe Gln Arg Ala Arg Gln Met Asn Ile Asp Val
178           500           505           510
180 Ala Thr Trp Val Arg Leu Leu Arg Arg Leu Ile Pro Asn Ala Thr Gly
181           515           520           525
183 Thr Gly Thr Phe Ser Pro Gly Ala Ser Pro Gly Ser Glu Ala Arg Thr
184           530           535           540
186 Thr Gly Asp Ile Ser Val Glu Lys Leu Asn Leu Gly Thr Asp Ser Asp
187 545           550           555           560
189 Ser Ser Pro Gln Lys Ser Ser Arg Asp Pro Pro Ser Ser Pro Ser Ser

```

RAW SEQUENCE LISTING

DATE: 08/20/2004

PATENT APPLICATION: US/10/817,418

TIME: 10:44:03

Input Set : A:\38002071.app

Output Set: N:\CRF4\08202004\J817418.raw

```

190          565          570          575
192 Leu Ser Ser Pro Ile Gln Glu Ser Thr Ala Pro Glu Leu Pro Ser Glu
193          580          585          590
195 Thr Gln Glu Thr Pro Gly Pro Ala Leu Cys Ser Pro Leu Arg Lys Ser
196          595          600          605
198 Pro Leu Thr Leu Glu Asp Phe Lys Phe Leu Ala Val Leu Gly Arg Gly
199          610          615          620
201 His Phe Gly Lys Val Leu Leu Ser Glu Phe Arg Pro Ser Gly Glu Leu
202 625          630          635          640
204 Phe Ala Ile Lys Ala Leu Lys Lys Gly Asp Ile Val Ala Arg Asp Glu
205          645          650          655
207 Val Glu Ser Leu Met Cys Glu Lys Arg Ile Leu Ala Ala Val Thr Ser
208          660          665          670
210 Ala Gly His Pro Phe Leu Val Asn Leu Phe Gly Cys Phe Gln Thr Pro
211          675          680          685
213 Glu His Val Cys Phe Val Met Glu Tyr Ser Ala Gly Gly Asp Leu Met
214          690          695          700
216 Leu His Ile His Ser Asp Val Phe Ser Glu Pro Arg Ala Ile Phe Tyr
217 705          710          715          720
219 Ser Ala Cys Val Val Leu Gly Leu Gln Phe Leu His Glu His Lys Ile
220          725          730          735
222 Val Tyr Arg Asp Leu Lys Leu Asp Asn Leu Leu Leu Asp Thr Glu Gly
223          740          745          750
225 Tyr Val Lys Ile Ala Asp Phe Gly Leu Cys Lys Glu Gly Met Gly Tyr
226          755          760          765
228 Gly Asp Arg Thr Ser Thr Phe Cys Gly Thr Pro Glu Phe Leu Ala Pro
229          770          775          780
231 Glu Val Leu Thr Asp Thr Ser Tyr Thr Arg Ala Val Asp Trp Trp Gly
232 785          790          795          800
234 Leu Gly Val Leu Leu Tyr Glu Met Leu Val Gly Glu Ser Pro Phe Pro
235          805          810          815
237 Gly Asp Asp Glu Glu Glu Val Phe Asp Ser Ile Val Asn Asp Glu Val
238          820          825          830
240 Arg Tyr Pro Arg Phe Leu Ser Ala Glu Ala Ile Gly Ile Met Arg Arg
241          835          840          845
243 Leu Leu Arg Arg Asn Pro Glu Arg Arg Leu Gly Ser Ser Glu Arg Asp
244          850          855          860
246 Ala Glu Asp Val Lys Lys Gln Pro Phe Phe Arg Thr Leu Gly Trp Glu
247 865          870          875          880
249 Ala Leu Leu Ala Arg Arg Leu Pro Pro Pro Phe Val Pro Thr Leu Ser
250          885          890          895
252 Gly Arg Thr Asp Val Ser Asn Phe Asp Glu Glu Phe Thr Gly Glu Ala
253          900          905          910
255 Pro Thr Leu Ser Pro Pro Arg Asp Ala Arg Pro Leu Thr Ala Ala Glu
256          915          920          925
258 Gln Ala Ala Phe Leu Asp Phe Asp Phe Val Ala Gly Gly Cys
259          930          935          940
262 <210> SEQ ID NO: 3
263 <211> LENGTH: 2829

```

RAW SEQUENCE LISTING

DATE: 08/20/2004

PATENT APPLICATION: US/10/817,418

TIME: 10:44:03

Input Set : A:\38002071.app

Output Set : N:\CRF4\08202004\J817418.raw

264 <212> TYPE: DNA

265 <213> ORGANISM: Homo sapiens

267 <400> SEQUENCE: 3

```

268 atggccagcg acgcegtgca gagtgagcct cgcagctggt ccctgctaga gcagctgggc 60
269 ctggccgggg cagacctggc ggcccccggg gtacagcagc agctggagct ggagcgggag 120
270 cggctgcggc gggaaatccg caaggagctg aagctgaagg aggtgctga gaacctgcgg 180
271 cgggccacca ctgacctggg ccgagcctg gggcccgtag agctgctgct gcggggctcc 240
272 tcgcccgcgc tcgacctgct gcaccagcag ctgcaggagc tgcacgcca cgtggtgctt 300
273 cccgaccccg cggccaccca cgatggcccc cagtccctg gtgcgggtgg cccacctgc 360
274 tcggccacca acctgagccg cgtggcgggc ctggagaagc agttggccat tgagctgaag 420
275 gtgaagcagg gggcgagaa catgatccag acctacagca atggcagcac caaggaccgg 480
276 aagctgctgc tgacagccca gcagatgttg caggacagta agaccaagat tgacatcatc 540
277 cgcattgcaac tccgcggggc gctgcaggcc gaccagctgg agaaccaggc agccccggat 600
278 gacaccaag ggagtctga cctgggggct gtggagctgc gcatcgaaga gctgcggcac 660
279 cacttccgag tggagcacgc ggtggccgag ggtgccaaga acgtactgcg cctgctcagc 720
280 gctgccaaag ccccgaccg caaggcagtc agcgaggccc aggagaaatt gacagaatcc 780
281 aaccagaagc tggggctgct gcgggaggct ctggagcgga gacttgggga gctgcccgc 840
282 gaccacccca agggcgggct gctgcgagaa gagctcgctg cggctcctc cgtgccttc 900
283 agcaccgcgc tggccgggccc ctttccgcgc acgcactaca gcacctgtg caagcccgcg 960
284 ccgtcacag ggacctgga ggtacgagtg gtgggtgca gagacctccc agagaccatc 1020
285 ccgtggaacc ctacccctc aatgggggga cctgggaccc cagacagccg ccccccttc 1080
286 ctgagccgccc cagcccgggg cctttacagc cgaagcggaa gcctcagtgg ccggagcagc 1140
287 ctcaaagcag aagccgagaa caccagtgaa gtcagcactg tgcttaagct ggataacaca 1200
288 gtggtggggc agacgtcttg gaagccatgt ggcccgaatg cctgggacca gagcttcaact 1260
289 ctggagctgg aaagggcacg ggaactggag ttggtgtgt tctggcggga ccagcggggc 1320
290 ctgtgtgccc tcaaatctc gaagttggag gatttcttg acaatgagag gcatgagggtg 1380
291 cagctggaca tggaaaccca gggctgcctg gtggctgagg tcaccttccg caacctgtc 1440
292 attgagagga ttctcggct ccgacggcag aagaaaattt tctccaagca gcaagggaag 1500
293 gcgttccagc gtgctaggca gatgaacatc gatgtcgcca cgtgggtgcg gctgctccgg 1560
294 aggtcatcc ccaatgccac gggcacaggc accttagcc ctggggcttc tccaggatcc 1620
295 gaggcccgga ccacgggtga catatcggtg gagaagctga acctcggcac tgactcggac 1680
296 agctcacctc agaagagctc gcgggatcct cttccagcc catcgacct gagctcccc 1740
297 atccaggaat ccactgctcc cgagctgctc tcggagaccc aggagacccc agggcccgcc 1800
298 ctgtgcagcc ctctgaggaa gtcacctctg acctcgaag atttcaagtt cctggcgggtg 1860
299 ctggccggg gtcatttttg gaaggtgtc ctctccgaat tccggcccag tggggagctg 1920
300 ttcgccatca aggtcttgaa gaaaggggac attgtggccc gagacgaggt ggagagcctg 1980
301 atgtgtgaga agcggatatt ggcggcagtg accagtgcg gacaccctt cctggtgaac 2040
302 ctcttcggct gtttccagac accggagcac gtgtgcttcg tgatggagta ctccggcgg 2100
303 ggggacctga tgctgcacat ccacagcgac gtgttctctg agccccgtgc catcttttat 2160
304 tccgctgcg tgggtgctgg cctacagttt cttcacgaac acaagatcgt ctacagggac 2220
305 ctgaagttgg acaatttgct cctggacacc gagggctacg tcaagatcgc agactttggc 2280
306 ctctgcaagg aggggatggg ctatggggac cggaccagca cattctgtgg gaccccgag 2340
307 ttcctggccc ctgaggtgct gacggacacg tcgtacacgc gagctgtgga ctggtgggga 2400
308 ctgggtgtgc tgctctacga gatgctggtt ggcgagtcct cattcccagg ggatgatgag 2460
309 gaggaggtct tcgacagcat cgtcaacgac gaggttcgct acccccgctt cctgtcggcc 2520
310 gaagccatcg gcatcatgag aaggctgctt cggaggaacc cagagcggag gctgggatct 2580
311 agcgagagag atgcagaaga tgtgaagaaa cagcccttct tcaggactct gggctgggaa 2640
312 gccctgttgg cccggcgcc gccaccgccc tttgtgccc cgtgtccgg ccgcaccgac 2700
313 gtcagcaact tcgacgagga gttcaccggg gaggccccc cactgagccc gccccgcgac 2760

```

VERIFICATION SUMMARY

DATE: 08/20/2004

PATENT APPLICATION: US/10/817,418

TIME: 10:44:04

Input Set : A:\38002071.app

Output Set: N:\CRF4\08202004\J817418.raw